# **DL - Lab 04**

# IT21251900 – Rajapaksha R.M.S.D

## **Question 7**

The reason there are 1,783 boxes is because the YOLO model splits the input image into a grid of 19x19 cells. Each cell predicts 5 bounding boxes, which gives a total of 1,805 boxes at the start (19x19x5 = 1,805). However, the yolo\_filter\_boxes function removes boxes with low confidence scores, leaving 1,783 boxes. Each of these boxes is described by four things: width, height, x-center, and y-center.

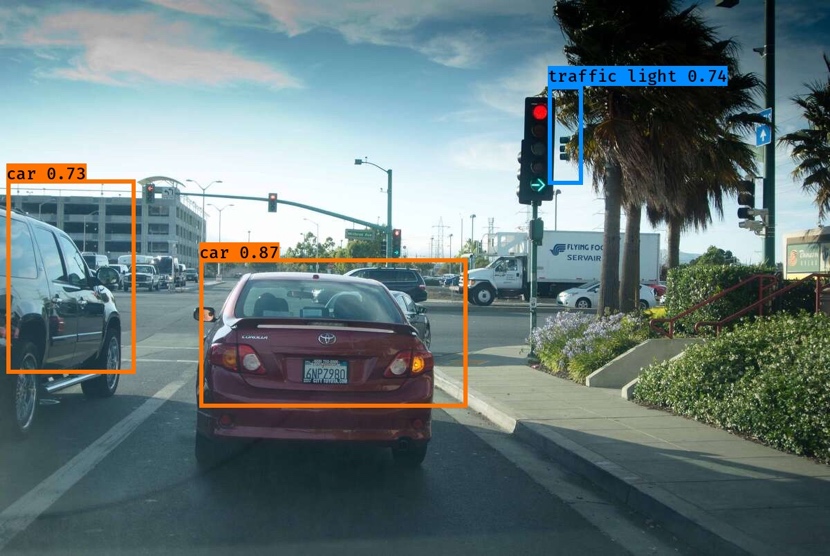
The highest number of boxes, 1,805, happens when all the predicted boxes have high confidence scores. The lowest number of boxes is 0, which occurs when none of the boxes pass the confidence threshold. So, after filtering, the number of boxes can be anywhere from 0 to 1,805.

## **Question 8**

Anchor boxes offer a big advantage by letting the model check all possible object predictions at once, instead of using a sliding window that requires separate predictions for each part of the image. This is efficient because anchor boxes have preset shapes and sizes, allowing the model to predict objects of different dimensions in a single pass through the image, which speeds up the detection process.

The sizes of anchor boxes are usually chosen through a combination of manual selection and data analysis. A common method is to use a clustering algorithm, like K-Means, on the dataset. This algorithm finds the most common shapes and sizes of bounding boxes in the dataset, which are then used as anchor boxes. By tailoring the anchor boxes to fit the specific data, the model can predict objects more accurately.

## **Question 09**



## **Question 10**



In this instance, no bounding boxes were detected despite the presence of vehicles and traffic lights. This lack of detection could be due to the unique shape of the vehicle, which might not have matched well with the anchor boxes or the model's predictions. Furthermore, the traffic lights might have gone unnoticed because of the lighting conditions in the image, which could have affected the model's ability to identify them.



In this image, both a vehicle and a traffic light were successfully detected. This observation supports the earlier assumptions, as the vehicle was detected even though only half of it was visible, indicating that the vehicle's shape or partial visibility in the previous image could have impacted the detection results. Additionally, the detection score for the traffic light in this image was very close to the threshold, which could explain why traffic lights in other images with less favorable conditions were not detected.

## **Question 11**

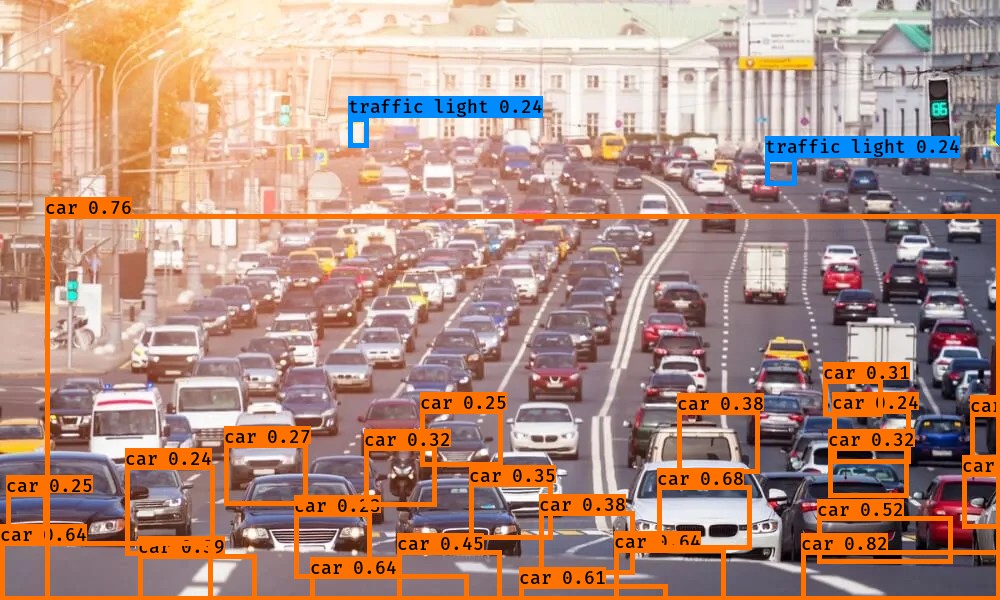
Default

**

max\_boxes = 40

score\_threshold = 0.2

iou\_threshold = 0.2



max\_boxes = 300

score\_threshold = 0.9

iou\_threshold = 0.9

**

Result: Detected none of them.

max\_boxes = 300

score\_threshold = 0.1

iou\_threshold = 0.9

A blurry image of a traffic jam

Description automatically generated

Result : Nothing is clear

max\_boxes = 300

score\_threshold = 0.1

iou\_threshold = 0.1

A traffic jam on a road

Description automatically generated

Result : detected many items correctly for a certain distance, can indentify back of the images compared to front